Linear Programming Duality

Vladimir Kolmogorov

1 Lagrangian duality

We start with an overview of Lagrangian duality for a general optimization problem:

\[ p^* := \min_{x \in X} p(x) \]  \hspace{1cm} (1a)

subject to \[ f_i(x) = 0 \quad \forall i \in I \]  \hspace{1cm} (1b)

\[ g_i(x) \leq 0 \quad \forall i \in I' \]  \hspace{1cm} (1c)

Usually we would need to write “inf” and “sup” instead of “min” and “max”. However, we will be interested in the case when problem (1) corresponds to a linear program; in this case min, max will suffice. We use the convention that the minimum over an empty set of arguments is +\(\infty\) (or \(-\infty\) for the maximum). Thus, if problem (1) does not have a feasible solution then \(p^* = +\infty\). If problem (1) is not bounded from below then \(p^* = -\infty\).

To form the Lagrangian dual problem, we introduce a dual variable for each constraint: vector \(\lambda \in \mathbb{R}^I\) for equality constraints (1b) and vector \(\mu \in \mathbb{R}^{I'}\) for inequality constraints (1c). Vector \(\mu\) will be required to be non-negative: \(\mu \geq 0\). We now define functions

\[ L(x, \lambda, \mu) = p(x) + \sum_{i \in I} \lambda_i f_i(x) + \sum_{i \in I'} \mu_i g_i(x) \quad \text{(Lagrangian)} \]  \hspace{1cm} (2)

\[ d(\lambda, \mu) = \min_{x \in X} L(x, \lambda, \mu) \quad \text{(dual function)} \]  \hspace{1cm} (3)

**Proposition 1.** For any \(\lambda, \mu\) with \(\mu \geq 0\) there holds \(d(\lambda, \mu) \leq p^*\).

**Proof.** Assume that \(p^* < +\infty\) (otherwise the claim is trivial). Let \(x^*\) be an optimal solution of (1) (or just a feasible solution if \(p^* = -\infty\)). We need to show that \(d(\lambda, \mu) \leq p(x^*)\).

\[
\begin{align*}
d(\lambda, \mu) &= \min_{x \in X} L(x, \lambda, \mu) \\
&\leq L(x^*, \lambda, \mu) = p(x^*) + \sum_{i=1}^m \lambda_i f_i(x^*) + \sum_{i=1}^{m'} \mu_i g_i(x^*) \\
&\leq p(x^*)
\end{align*}
\]

We showed that \(d(\lambda, \mu)\) is a lower bound on \(p^*\). To obtain the tightest bound, we formulate the following maximization problem:

\[ d^* := \max_{\lambda, \mu} d(\lambda, \mu) \]  \hspace{1cm} (4a)

subject to \(\mu \geq 0\)  \hspace{1cm} (4b)

By Proposition 1, \(d^* \leq p^*\). This relation is known as weak duality. The difference \(p^* - d^* \geq 0\) is called the duality gap.
**Strong duality** If \( d^* = p^* \) then problems (1) and (4) are said to have a *strong duality*. This is often the case when problem (1) is *convex*.

**Definition 2.** Problem (1) is called convex if (i) set \( X \) is convex; (ii) functions \( p \) and \( g_i \) are convex; (iii) functions \( f_i \) are linear.

In general, convexity of (1) does not guarantee a strong duality. However, an additional mild condition known as the *(weak) Slater condition* will ensure that \( d^* = p^* \). For linear programs this condition reduces to just feasibility.

**Theorem 3.** Suppose that functions \( p, f_i, g_i : \mathbb{R}^n \to \mathbb{R} \) are linear and set \( X \) is a polyhedron, i.e. \( X = \{ x \in \mathbb{R}^n \mid Ax \leq b \} \) for some matrix \( A \in \mathbb{R}^{m \times n} \) and vector \( b \in \mathbb{R}^{m \times 1} \). If at least one of the values \( p^*, d^* \) is finite then \( d^* = p^* \).

### 2 Duality for linear programs

We now apply the Lagrangian approach to linear programs. We will illustrate the derivation on a numerical example with two variables; the derivation in the general case uses exactly the same ideas.

#### 2.1 Example with two variables

Consider the following problem.

\[
p^* := \min_{x,y : y \geq 0} \quad 2x + 3y \tag{5a}
\]

subject to \( x + 2y = 4 \) \quad (5b)

\( 3x - y \geq 5 \) \quad (5c)

The Lagrangian function is given by

\[
L(x, y, \lambda, \mu) = (2x + 3y) + \lambda \cdot [4 - (x + 2y)] + \mu \cdot [5 - (3x - y)]
\]

\[
= (4\lambda + 5\mu) + [2 - (\lambda + 3\mu)] \cdot x + [3 - (2\lambda - \mu)] \cdot y
\]

From \( L \) we obtain the dual function:

\[
d(\lambda, \mu) = \min_{x,y : y \geq 0} L(x, y, \lambda, \mu) = \begin{cases} 
4\lambda + 5\mu & \text{if } 2 - (\lambda + 3\mu) = 0 \text{ and } 3 - (2\lambda - \mu) \geq 0 \\
-\infty & \text{otherwise}
\end{cases}
\]

Therefore, the dual problem can be written as

\[
d^* := \max_{\lambda,\mu : \mu \geq 0} \quad 4\lambda + 5\mu \tag{6a}
\]

subject to \( \lambda + 3\mu = 2 \) \quad (6b)

\( 2\lambda - \mu \leq 3 \) \quad (6c)

#### 2.2 General LPs

Consider two linear programs marked as “PRIMAL” and “DUAL”:

<table>
<thead>
<tr>
<th>PRIMAL</th>
<th>DUAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \min c^\top x )</td>
<td>( \max b^\top y )</td>
</tr>
<tr>
<td>( a_i x = b_i \quad i \in I )</td>
<td>( y_i \geq 0 )</td>
</tr>
<tr>
<td>( a_i x \geq b_i \quad i \in I' )</td>
<td>( y_i \geq 0 )</td>
</tr>
<tr>
<td>( x_j \geq 0 \quad j \in J )</td>
<td>( a_j^\top y = c_j )</td>
</tr>
<tr>
<td>( x_j \geq 0 \quad j \in J' )</td>
<td>( a_j^\top y \leq c_j )</td>
</tr>
</tbody>
</table>
Here \( A = (a_{ij}) \) is a matrix of size \( m \times n \), and \( c \in \mathbb{R}^n \), \( b \in \mathbb{R}^m \) are column vectors. The \( i \)-th row of \( A \) is denoted as \( a_i \), and the \( j \)-th row of \( A^\top = (a_{ji}) \) (or equivalently the transposed \( j \)-th column of \( A \)) is denoted as \( a_j^\top \). Note that \( a_i \) and \( a_j^\top \) are row vectors. Problem \textsc{primal} has

- \( n = |J| + |J'| \) variables; set \( J \) indexes unrestricted variables and set \( J' \) indexes non-negative variables.
- \( m = |I| + |I'| \) constraints; set \( I \) indexes equality constraints and set \( I' \) indexes inequality constraints.

In \textsc{dual} the variables and constraints are swapped. The optimal values of \textsc{primal} and \textsc{dual} will be denoted as \( p^* \) and \( d^* \) respectively.

It can be shown that applying the Lagrangian approach to \textsc{primal} gives the problem \textsc{dual}, and vice versa. The derivations of these facts are very similar to that in section 2.1. We will omit these derivations; instead, we will re-establish the weak duality relation, and then use it for deriving complementary slackness conditions.

We say that \((x, y)\) is a feasible primal-dual pair if \( x \) and \( y \) satisfy constraints of problems \textsc{primal} and \textsc{dual} respectively. If in addition \( b^\top y = c^\top x \) then it is called optimal.

The following proposition implies that \( d^* \leq p^* \).

**Proposition 4** (Weak duality). If \((x, y)\) is a feasible primal-dual pair then \( b^\top y \leq c^\top x \).

*Proof.* We can write

\[
b^\top y \leq (Ax)^\top y = x^\top (A^\top y) \leq x^\top c = c^\top x
\]  

(8)

Let us prove the first inequality (the proof of the second one is similar):

\[
b^\top y - (Ax)^\top y = \sum_{i=1}^{n} \underbrace{[b_i - (a_i x)] \cdot y_i}_{\leq 0} \leq 0
\]  

(9)

(To show that each term in the second sum is non-positive we need to consider separately cases \( i \in I \) and \( i \in I' \).)

**Proposition 5** (Complementary slackness). Let \((x^*, y^*)\) be a feasible primal-dual pair. It is optimal if and only if the following conditions hold:

\[
\begin{align*}
[a_i x - b_i] \cdot y_i &= 0 & \forall i \in I \cup I' \\
[a_j^\top y - c_j] \cdot x_j &= 0 & \forall j \in J \cup J'
\end{align*}
\]  

(10a)

(10b)

Equivalently, dual variables corresponding to non-tight constraints must be zero:

\[
\begin{align*}
a_i x > b_i &\implies y_i = 0 & \forall i \in I' \\
a_j^\top y < c_j &\implies x_j = 0 & \forall j \in J'
\end{align*}
\]  

(11a)

(11b)

*Proof.* If conditions (10) are satisfied then all inequalities in (9) are tight, and thus so is inequality (i) in (8). Similarly, we prove that inequality (ii) in (8) is also tight. Therefore, \( b^\top y = c^\top x \).

Conversely, suppose that \( b^\top y = c^\top x \); then inequalities (i) and (ii) in (8) are tight. Therefore, the expression in (8) is zero. Since terms in (8) are non-positive, each term must be zero, i.e. (10a) holds. The proof of (10b) is similar.
**Strong duality and infeasible problems**  The following result is fundamental in the theory of Linear Programming.

**Theorem 6.** *(Strong duality)* If at least one of the one of the values $p^*$, $d^*$ is finite then $p^* = d^*$.

If we consider infeasible problems then the relation $p^* = d^*$ may not hold: we may have $(p^*, d^*) = (\infty, -\infty)$ (i.e. both problems are infeasible). Possible situations are summarized in the table below (cases marked with $\times$ can be eliminated by the weak duality).

<table>
<thead>
<tr>
<th>primal</th>
<th>finite optimum</th>
<th>unbounded</th>
<th>infeasible</th>
</tr>
</thead>
<tbody>
<tr>
<td>finite optimum</td>
<td>ok</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>unbounded</td>
<td>$\times$</td>
<td>$\times$</td>
<td>ok</td>
</tr>
<tr>
<td>infeasible</td>
<td>$\times$</td>
<td>ok</td>
<td>ok</td>
</tr>
</tbody>
</table>